Amendments to the Claims

Claims 12 and 59 are amended. Claims 12-24 and 59 are currently pending.

1. - 11. (Cancelled)

12. (Presently Amended) A method for scoring a match of two peptides, the method comprising:

providing information associated with an experimental peptide, where the information comprises at least mass spectrum information associated with the experimental peptide and at least one fragment of the experimental peptide;

providing information associated with a candidate peptide;

defining an extended match E based on the information associated with the experimental peptide and the information associated with the candidate peptide, the extended match E being a probabilistic function of a tuple of random variables that include at least a fragment match and a peptide match between the experimental peptide and the candidate peptide;

generating a stochastic model based on the information associated with the experimental peptide and the information associated with the candidate peptide, the stochastic model incorporating a probability distribution of each of the random variables;

scoring the extended match E based on a likelihood ratio $L = \frac{\mathbf{P}(E|D,s,H_1)}{\mathbf{P}(E|D,s,H_0)}$, where

D is any extra information that is associated with the experimental peptide and the candidate peptide;

s is a peptide sequence associated with the candidate peptide;

H_1 is a hypothesis that the peptide sequence s is the correct sequence of the experimental peptide;

H_0 is a null-hypothesis that the peptide sequence s is an erroneous sequence of the experimental peptide; and

probabilities $\mathbf{P}(E|D,s,H_1)$ and $\mathbf{P}(E|D,s,H_0)$ are calculated based on the stochastic model; and

outputting, in a user readable format, a result of ~~information based at least in part on~~ the step of scoring the extended match E.

13. (Previously Amended) The method according to claim 12, where the extended match E is a random variable that further comprises one or more random variables, the one or more random variables being selected from a group consisting of:

peptide match P that characterizes a match between the experimental peptide mass m and the candidate peptide mass m_i;

fragment match F that characterizes a match between fragment masses f_j of the experimental peptide and fragment masses $m_{i,j}$ of the candidate peptide, where j is an index for the fragment masses of the experimental peptide;

charge z that is used to match the m/z ratio of the experimental peptide with the candidate peptide;

elution time t of the experimental peptide;

number of missed cleavages k in the candidate peptide matching the experimental peptide;

protein/peptide modifications W made to the candidate peptide to match the experimental peptide; and

any random variables observable or derivable based on the information associated with the experimental peptide and the candidate peptide.

14. (Previously Amended) The method according to claim 13 further comprising determining the probability distributions for the one or more random variables based on at least one of a Hidden Markov Model and an artificial neural network.

15. (Previously Amended) The method according to claim 13 further comprising determining an empirical probability distribution for the one or more random variables based on matches between experimental data for known peptides and peptides in a peptide database.

16. (Original) The method according to claim 12 further comprising estimating the probabilities $\mathbf{P}(E|D,s,H_1)$ and $\mathbf{P}(E|D,s,H_0)$ based on the lemma $\mathbf{P}(A,B|C) = \mathbf{P}(A|B,C)\ \mathbf{P}(B|C)$, where A, B and C are random variables.

17. (Original) The method according to claim 12 further comprising calculating at least one of:

Bayesian score $L' = \dfrac{\mathbf{P}(H_1|D,s,E)}{\mathbf{P}(H_0|D,s,E)} = L\dfrac{\mathbf{P}(H_1|D,s)}{\mathbf{P}(H_0|D,s)}$; and

Bayesian score $L'' = L\dfrac{\mathbf{P}(H_1|D,s,Q)}{\mathbf{P}(H_0|D,s,Q)}$, where Q represents statistics associated with mass spectrum quality of the experimental peptide.

18. (Original) The method according to claim 12 further comprising:

comparing the candidate peptide mass with the experimental peptide mass; and

scoring the extended match E based on the likelihood ratio L, if the difference between the candidate peptide mass and the experimental peptide mass is in a predetermined range.

19. (Original) The method according to claim 12 further comprising adjusting the stochastic model and a plurality of parameters associated with the stochastic model based on a learning data set, where the learning data set comprises a plurality of peptides that have been identified or a set of known protein standards.

20. (Original) The method according to claim 12 further comprising generating an output, where the output comprises at least one of:

a match score associated with the candidate peptide, where the match score comprises at least one of

the likelihood;

a log-likelihood, where the log-likelihood is the logarithm of the likelihood ratio;

the likelihood ratio divided by the length of the experimental peptide measured in amino acids;

the log-likelihood divided by the length of the experimental peptide measured in amino acids; and

the log-likelihood divided by the logarithm of the length of the experimental peptide measured in amino acids;

a Z-score associated with the match score;

a p-value associated with the match score;

biological information associated with the experimental peptide; and

biological information associated with the candidate peptide.

21. (Original) The method according to claim 12, where a theoretical fragmentation spectrum is provided for the candidate peptide.

22. (Original) The method according to claim 21, where the theoretical fragmentation spectrum includes masses corresponding to fragment isotopes.

23. (Original) The method according to claim 12 further comprising filtering the candidate peptide based on at least one of:

the taxonomy of the protein that the candidate peptide belongs to;

the isoelectric point of the protein that the candidate peptide belongs to;

the molecular weight of the protein that the candidate peptide belongs to;

a non-symmetric mass window; and

a set of possible masses made of the union of a plurality of mass intervals.

24. (Original) The method according to claim 12 further comprising

providing a physical sample of the experimental peptide and biological information associated with the experimental peptide; and

providing a physical sample of the candidate peptide and biological information associated with the candidate peptide.

25. - 58. (Cancelled)

59. (Presently Amended) A method for scoring a match of two peptides, the method comprising:

providing information associated with an experimental peptide, where the information comprises at least mass spectrum information associated with the experimental peptide and at least one fragment of the experimental peptide;

providing information associated with a candidate peptide;

defining an extended match E based on the information associated with the experimental peptide and the information associated with the candidate peptide, the extended match E being a probabilistic function of a tuple of random variables that include at least a consecutive fragment match between the experimental peptide and the candidate peptide;

generating a stochastic model based on the information associated with the experimental peptide and the information associated with the candidate peptide, the stochastic model incorporating a probability distribution of each of the random variables, wherein the probability distribution associated with the consecutive fragment match is determined based on a Hidden Markov Model;

scoring the extended match E based on a likelihood ratio $L = \frac{\mathbf{P}(E|D,s,H_1)}{\mathbf{P}(E|D,s,H_0)}$, wherein:

D is any extra information that is associated with the experimental peptide and the candidate peptide,

s is a peptide sequence associated with the candidate peptide,

H_1 is a hypothesis that the peptide sequence s is the correct sequence of the experimental peptide,

H_0 is a null-hypothesis that the peptide sequence s is an erroneous sequence of the experimental peptide,

probabilities $\mathbf{P}(E|D,s,H_1)$ and $\mathbf{P}(E|D,s,H_0)$ are calculated based on the stochastic model, and

the scoring further takes into account a charge state match between the experimental peptide and the candidate peptide; and

outputting, in a user readable format, a result of ~~information based at least in part on~~ the step of scoring the extended match E.